Statistica Sinica Preprint No: SS-2024-0224

Title	General Sliced Factorial Designs for Online Experiments with Multiple Platforms
Manuscript ID	SS-2024-0224
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0224
Complete List of Authors	Zijian Han, Dongying Wang, Fasheng Sun and Peter Chien
Corresponding Authors	Peter Chien
E-mails	peter.chien@wisc.edu
Notice: Accepted author version.	

Statistica Sinica

General Sliced Factorial Designs for Online Experiments with Multiple Platforms

Zijian Han[1], Dongying Wang[2], Fasheng Sun[1] and Peter Chien[3,*]

[1]*Northeast Normal University,* [2]*Jilin University of Finance and Economics*

[3]*University of Wisconsin-Madison*

Abstract: Digital marketing is an integral part of digital transformation in industry. It is critical to use design of experiments to conduct online experiments for various forms of digital marketing including web design, email marketing, social network and others. Online experiments often involve multiple platforms, including desktop computers from different manufacturers, different types of mobiles and smart watches of different brands. A sliced factorial design is a suitable choice for designing online experiments with multiple platforms. We provide a general theory for sliced factorial designs and propose sliced generalized wordlength patterns to construct such designs for any number of platforms. The theory uses the characteristics of parallel flat design-based sliced factorial designs to construct optimal sliced factorial designs. For practical use, several constructed designs are provided in the Supplementary Material, enabling practitioners to efficiently design and analyze online experiments across multiple platforms.

Key words and phrases: Design of experiments, J-characteristics, Digital marketing, Sliced effect hierarchy principle.

1. Introduction

With the increasing growth of the internet, online experiments are being conducted for various forms of digital marketing including website optimization, mobile apps, social networks, video recommendation and email campaigns (Haizler and Steinberg, 2021).

Typical A/B testing refers to experiments whose goal is to identify whether a proposed change is effective. The term "A/B testing" was invented for online experiments in which a sample population is randomly divided into two mutually exclusive sets (runs), A and B, allowing for comparative analysis. Some online experiments test features one-at-a-time, which is called "one-factor-at-a-time" (OFAT) in the design of experiments. Since the pioneering work of Fisher, it has been well known that factorial experiments outperform OFAT (see section 4.7 in Wu and Hamada (2021)). This form of testing is called a multivariate test in online experiments. In a simple example, consider the following three website factors, each at two levels, corresponding to two different versions: the "Navigation" tab on top left, the "Login" tab on top right, and the "Research" tab in the middle.

To find the best combination of these three attributes, one can create eight versions of the website for a full factorial design. To avoid confusion, we define "run size" as the number of distinct level combinations, and "sample size" as the number of users assigned to the level combinations. As the number of factors increases, it is necessary in online experiments to use a fraction of all possible combinations to conduct multivariate tests (Wu and Hamada, 2021) (Haizler and Steinberg, 2021). Fractional factorial designs as such are particularly advantageous in the online space for the following reasons:

1. **Cost Considerations**: The setup and execution costs for each factor level combination are significant and increase as the number of factors increases.

2. **High-dimensional Inputs**: In online experiments with high-dimensional input spaces, higher-order interaction effects are often negligible. This makes fractional factorial design both a practical and cost-effective choice. By reducing the number of combinations tested, researchers can allocate larger sample sizes to each combination, improving inference accuracy.

From an experimental design point of view, a new challenge in the digital space is that online experiments are often conducted across multiple platforms including different electronic devices and different operational systems. For example, Sadeghi et al. (2020) used a multi-platform experiment for an empirical email optimization application to maximize engagement for a digital magazine. They proposed sliced factorial designs for multi-platform experiments. The slicing idea in such a design came from sliced space-filling designs in computer experiments where the focus is on low-dimensional projections. References in this direction include Qian and Wu (2009), Qian (2012), Yang et al. (2014), and Kong et al. (2018), among others. Sadeghi et al. (2020) introduced the idea of sliced wordlength pattern to construct sliced factorial designs (SFDs) with a factorial structure. Guo et al. (2024) extended such designs to handle multiple layers with two platforms and Sadeghi et al. (2024) constructed these designs for an industrial email campaign with four platforms.

However, these existing SFDs are only for two or four platforms and do not provide an explicit expression of the sliced wordlength pattern. As shown in Figure 1, many modern online experiments are conducted across many devices. For example, responsive web design or responsive design is an approach to web design that aims to make web pages render well on a variety of devices and window or screen sizes from minimum to maximum

Figure 1: Web site tested across many devices: https://www.smashingmagazine.com/2014/07/testing-and-responsive-web-design/.

display size to ensure usability and satisfaction (Wikipedia). A quick search on Amazon or BestBuy would show many smartphones, laptops, notebooks, smart watches and other electric devices for internet use. To address this practical challenge, we propose a method for constructing new SFDs for any number of platforms. This method views a sliced factorial design with n design factors and a slice factor as an asymmetrical fractional factorial design through the lens of the generalized minimum aberration (GMA, Xu and Wu 2001), which provides a new sliced generalized wordlength pattern for SFDs for multi-platform experiments.

We construct SFDs via parallel flats design (PFD) to accommodate any number of platforms. Connor and Young (1961) first proposed the idea of PFD, in which each single flat is a regular fractional factorial design

(FFD) and the complete design is a nonregular design that retains some properties of regular fractional factorial designs. For more details of PFDs, see Section 15.12 in Cheng (2014) and references therein. Recently, Edwards and Mee (2023) proposed a Kronecker product construction for nonregular designs, which essentially refers to PFDs. Wang and Mee (2021) provided a comprehensive review of two-level PFDs and developed a general theory.

The remainder of the article is organized as follows. Section 2 introduces the generalized minimum aberration criterion and design criteria for selecting factorial designs. Section 3 proposes the sliced generalized wordlength pattern and the sliced generalized minimum aberration criterion on the basis of the sliced effect hierarchy principle. Section 4 presents theoretical results on the characteristics of PFD-based sliced factorial designs and the methods for constructing optimal sliced factorial designs. Section 5 discusses the proposed designs with subdesigns of different run sizes. Section 6 concludes the article with a discussion and suggestions for future work. Some optimal SFDs are tabulated in the Supplementary Material for practical use.

2. Notation and definitions

An asymmetrical design of N runs, n factors and levels $s_1, \ldots, s_n$ is denoted the $(N, s_1 \cdots s_n)$-design. An $(N, s_1 \cdots s_n)$-design D is a set of N row vectors or an $N \times n$ matrix in which each row represents a run and each column represents a factor. The jth column of D takes values from a set of s_j symbols, e.g., $\{0, 1, \ldots, s_j - 1\}$ denoted as $\mathbf{Z}_{s_j}$, which is the integer ring with modulo s_j, and each row of D is a point from $H = \mathbf{Z}_{s_1} \times \cdots \times \mathbf{Z}_{s_n}$.

Consider the general ANOVA model $E(Y(\boldsymbol{x})) = \sum_{\boldsymbol{u} \in H} \chi_{\boldsymbol{u}}(\boldsymbol{x}) \beta_{\boldsymbol{u}}$, where $Y(\boldsymbol{x})$ is the response of the design point $\boldsymbol{x} \in H$, $\beta_{\boldsymbol{u}}$s are factorial effects, and $\{\chi_{\boldsymbol{u}}, \boldsymbol{u} \in H\}$ are orthonormal contrast coefficients such that

$$\sum_{\boldsymbol{x} \in H} \chi_{\boldsymbol{u}}(\boldsymbol{x}) \overline{\chi_{\boldsymbol{v}}(\boldsymbol{x})} = |H| \delta_{\boldsymbol{u},\boldsymbol{v}},$$

where $\overline{\chi_{\boldsymbol{v}}(\cdot)}$ is the complex conjugate of $\chi_{\boldsymbol{v}}(\cdot)$. $|H|$ is the number of elements in the set H, and $\delta_{\boldsymbol{u},\boldsymbol{v}}$ is the Kronecker delta function, which equals 1 if $\boldsymbol{u} = \boldsymbol{v}$ and 0 otherwise. As in Wu and Xu (2001), only contrasts defined by tensor products are considered:

$$\chi_{\boldsymbol{u}}(\boldsymbol{x}) = \prod_{i=1}^{n} \chi_{u_i}^{(s_i)}(x_i) \quad \text{for } \boldsymbol{u} = (u_1, \ldots, u_n) \in H, \quad \boldsymbol{x} = (x_1, \ldots, x_n) \in H,$$

where $\{\chi_{u_i}^{(s_i)}, u_i \in \mathbf{Z}_{s_i}\}$ are orthonormal contrasts for the ith factor with s_i

levels satisfying

$$\sum_{x_i \in \mathbf{Z}_{s_i}} \chi_{u_i}^{(s_i)}(x_i)\overline{\chi_{v_i}^{(s_i)}(x_i)} = s_i \delta_{u_i, v_i}$$

for any $u_i, v_i \in \mathbf{Z}_{s_i}$. Following Wu and Xu (2001),

$$A_j(D) = N^{-2} \sum_{wt(\boldsymbol{u})=j} |\chi_{\boldsymbol{u}}(D)|^2 \quad \text{for } j = 1, \ldots, n \tag{2.1}$$

where the summation is over all $\boldsymbol{u} \in H$ with j nonzero elements and $\chi_{\boldsymbol{u}}(D) = \sum_{\boldsymbol{x} \in D} \chi_{\boldsymbol{u}}(\boldsymbol{x})$. The vector $(A_1(D), \ldots, A_n(D))$ is called the generalized wordlength pattern, which results in the GMA criterion by sequentially minimizing $A_j(D)$ in (2.1) for $j = 1, \ldots, n$.

Designs with two-level factors are widely used in practice. We now focus on SFDs with two-level subdesigns conducted on each platform. That is $s_1 = \cdots = s_n = 2$. For the special case with $s_1 = \cdots = s_n = 2$, D is considered an $(N, 2^n)$-design, and $A_j(D)$ in (2.1) has the following equivalent representation:

$$A_j(D) = N^{-2} \sum_{wt(\boldsymbol{u})=j} J_{\boldsymbol{u}}(D)^2 \quad \text{for } j = 1, \ldots, n, \tag{2.2}$$

where the summation is over all n-tuple binary vectors $\boldsymbol{u}$ with j nonzero elements and $J_{\boldsymbol{u}}(D) = \sum_{\boldsymbol{x} \in D} (-1)^{\langle \boldsymbol{u}, \boldsymbol{x} \rangle}$ with $\langle \boldsymbol{u}, \boldsymbol{x} \rangle = \sum_{i=1}^n u_i x_i$. Tang and Deng (1999) defined $J_{\boldsymbol{u}}(D)$'s as J-characteristics of D and proposed the minimum G_2-aberration ($\mathrm{MG_2A}$) criterion to sequentially minimize $A_j(D)$ in (2.2) for $j = 1, \ldots, n$. GMA is the generalized version of $\mathrm{MG_2A}$.

3. A new sliced generalized wordlength pattern

Consider the sliced factorial design for a multi-platform experiment conducted on s platforms. Assume that the number of runs for the s platforms is $N_0, N_1, \ldots, N_{s-1}$, and for each platform, the number of design factors and the corresponding level setting are the same, resulting in n design factors with two levels. The sliced factorial designs are defined as follows.

Definition 1. We consider a multi-platform experiment for studying n design factors, denoted $F_1, \ldots, F_n$, on s platforms $P_0, \ldots, P_{s-1}$. Suppose that the design factor F_j has two levels, $j = 1, \ldots, n$, and that the slice factor S has s levels, with the level i being associated with P_i, $i \in \mathbf{Z}_s$. Let d_i be an $(N_i, 2^n)$-design conducted on P_i, $i \in \mathbf{Z}_s$. The sliced factorial design (SFD) is then defined as an $(\sum_{i=0}^{s-1} N_i, 2^n s)$-design D that contains s subdesigns $d_0, \ldots, d_{s-1}$ corresponding to the s levels of the slice factor S. In particular, if $N_i = N$ for $i = 0, \ldots, s-1$, SFD is an $(sN, 2^n s)$-design.

Online experiments are used for treatment comparisons, variable screening, system optimization and other purposes. The multi-platform experiments defined above differ from the experiments with blocking as follows. For the latter, the block factor is assumed to have no interaction with the treatment factor (Zhang and Park (2000)). In the former, the interactions between the slice factor and a design factor are not only nonnegligible but also important. Thus, multi-platform experiments are much more complex than experiments with blocking. In addition, because of the need to run across multiple platforms, the optimal level combination of the design factors often varies with the platform. It is important for sliced factorial designs to screen out the active design factors and accurately estimate the design factor effects in addition to the slice factor effect and interactions of the design factors and the slice factor. Hereinafter, all the effects involving the slice factor are referred to as slice factor effects. To be more precise, we present the following sliced effect hierarchy principle from Sadeghi et al. (2020):

(i) *All the lower-order effects are more likely to be important than the higher-order effects are;*

(ii) *For the slice factor or any design factor, effects of the same order are equally likely to be important;*

(iii) *Any slice factor effect is likely to be more important than is a design factor effect that is of the same order.*

To thoroughly examine the confounding structure of a sliced factorial design, we propose the following new sliced generalized wordlength pattern to measure the overall aliasing between all effects and the grand mean.

For $j = 1, \ldots, n$, define

$$A_{j,0}(D) = \frac{1}{(\sum_{i=0}^{s-1} N_i)^2} \sum_{wt(\boldsymbol{u})=j} \left(\sum_{i=0}^{s-1} J_{\boldsymbol{u}}(d_i) \right)^2, \tag{3.3}$$

$$A_{j,1}(D) = \frac{1}{(\sum_{i=0}^{s-1} N_i)^2} \sum_{v=1}^{s-1} \sum_{wt(\boldsymbol{u})=j-1} \left(\sum_{i=0}^{s-1} J_{\boldsymbol{u}}(d_i) \chi_v^{(s)}(i) \right)^2. \tag{3.4}$$

where $J_{\boldsymbol{u}}(d_i)$ and $\boldsymbol{u} \in \mathbf{Z}_2^n$ are the J-characteristics of subdesign d_i and where $\chi_v^{(s)}(\cdot)$ is the orthonormal contrast for the slice factor. The value $A_{j,0}$ measures the overall aliasing between all j-factor interactions $F_{i_1} \times \cdots \times F_{i_j}$ and the general mean, whereas $A_{j,1}$ measures the overall aliasing between all j-factor interactions $F_{i_1} \times \cdots \times F_{i_{j-1}} \times S$ and the general mean. According to the sliced effect hierarchy principle defined in Sadeghi et al. (2020), $A_{j+1,1}$ is less important than is $A_{j,0}$ but more important than is $A_{j+1,0}$. Thus, we define a sliced generalized wordlength pattern and the corresponding sliced generalized minimum aberration criterion.

Definition 2. For the sliced factorial $(\sum_{i=0}^{s-1} N_i, 2^n s)$-design D, the vector

$$(A_{1,1}, A_{1,0}, A_{2,1}, \ldots, A_{n,0}, A_{n+1,1}) \tag{3.5}$$

Is defined as the sliced generalized wordlength pattern (SGWLP). Then, the sliced generalized minimum aberration (SGMA) criterion sequentially minimizes each term in (3.5).

We provide one example to illustrate the definition of the SGWLP.

Example 1. For a multi-platform experiment with eight two-level design factors and a five-level slice factor, a $(5 \cdot 16, 2^8 5)$-design D, which consists of five subdesigns $d_0, \ldots, d_4$, is appropriate. Two SFDs D_1 and D_2 are constructed. D_1 contains the same subdesigns as $d_i = d_0$, $i = 1, 2, 3, 4$ contains, and d_0 is 2^{8-4} regular MA designs with generators $E = ABC, F = ABD, G = ACD$ and $H = BCD$. Design D_2 with five subdesigns defined as follows contains the same d_0 that D_1 contains, but d_1, d_2, d_3 and d_4 are isomorphic designs of d_0.

$$
\begin{array}{lllll}
d_0: A\ B\ C\ D\ E = ABC & F = ABD & G = ACD & H = BCD \\
d_1: A\ B\ C\ D\ E = ABC & F = ABD & G = ACD + \mathbf{1} & H = BCD + \mathbf{1} \\
d_2: A\ B\ C\ D\ E = ABC & F = ABD + \mathbf{1} & G = ACD & H = BCD + \mathbf{1} \\
d_3: A\ B\ C\ D\ E = ABC + \mathbf{1} & F = ABD & G = ACD + \mathbf{1} & H = BCD \\
d_4: A\ B\ C\ D\ E = ABC + \mathbf{1} & F = ABD + \mathbf{1} & G = ACD & H = BCD
\end{array}
$$

The addition operation is performed over $GF(2)$, and the $+\mathbf{1}$ operation switches elements 0 and 1 in the corresponding columns.

The SGWLPs of D_1 and D_2 are $(0^7, 14, 0^7, 1, 0)$ and $(0^7, 1.2, 12.8, 0^6, 1, 0)$, and 0^t is used to represent t successive zero components in SGWLP hereafter. Thus, D_2 is better than is D_1 on the basis of the SGMA criterion.

The design in this example can be used for digital marketing across five different mobiles, such as the top five mobiles in Global, Apple, Samsung, Xiaomi, OPPO and TRANSSION, as shown in Figure 2.

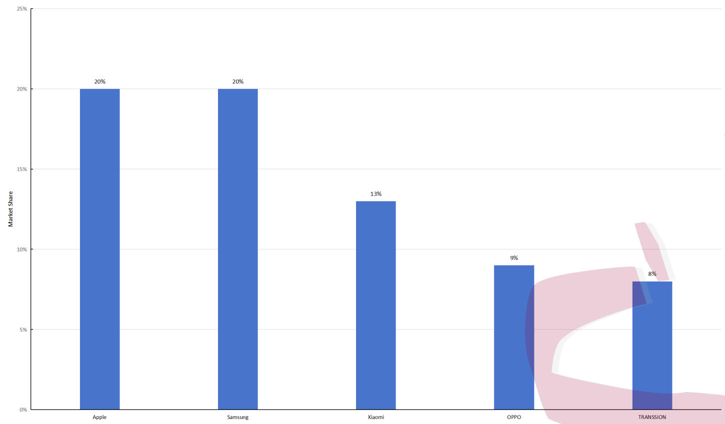

Figure 2: Top five mobiles in Global: https://www.canalys.com/newsroom/worldwide-smartphone-market-2023.

4. PFD-based sliced factorial design

In fact, D_2 in Example 1 is a PFD if the slice factor is deleted. This example implies the possibility of using the concise structure of PFDs to construct suitable SFDs under the SGMA criterion. Additionally, because the number of all possible candidates of SFDs for some specific parameters is very large, an exhaustive search is not feasible. We focus on constructing PFD-based sliced factorial designs.

4.1 Characteristics of the PFD-based sliced factorial design

We study the characteristics of the SGWLP for PFD-based SFDs to simplify the SGWLP and facilitate the construction of SGMA $(sN, 2^n s)$-designs. All the designs d_i are isomorphic via level permutation on columns and thus $N_i = N$ for $i = 0, \ldots, s-1$.

For convenience, a switch matrix $P = (p_{ij})_{s\times n}$ with $p_{ij} = 0, 1$ for $i = 0, \ldots, s-1$, where $j = 1, \ldots, n$. Without loss of generality, let $p_{0j} = 0$, where $j = 1, \ldots, n$. Then, the jth column of subdesign d_i is defined as

$$\boldsymbol{d}_{ij} = p_{ij} \oplus \boldsymbol{d}_{0j}, \text{ for } i = 0, \ldots, s-1, \; j = 1, \ldots, n. \tag{4.6}$$

Furthermore, $J_{\boldsymbol{u}}(\boldsymbol{p}_i) = (-1)^{\langle \boldsymbol{u}, \boldsymbol{p}_i \rangle}$ and $J_{\boldsymbol{u}}(P) = \sum_{i=0}^{s-1} J_{\boldsymbol{u}}(\boldsymbol{p}_i)$, where $\boldsymbol{p}_i$ is the $(i+1)$th row of the switch matrix P, and the J-characteristic of subdesign d_i can be expressed as

$$J_{\boldsymbol{u}}(d_i) = J_{\boldsymbol{u}}(d_0) J_{\boldsymbol{u}}(\boldsymbol{p}_i), \quad i = 0, \ldots, s-1. \tag{4.7}$$

The following theorem shows that the sliced generalized wordlength pattern of D is determined by the initial design d_0 and the switch matrix P.

Theorem 1. *For an $(sN, 2^n s)$-design consisting of s subdesigns $d_0, \ldots, d_{s-1}$*

defined in (4.6),

$$A_{j,0}(D) = \frac{1}{(sN)^2} \sum_{wt(\boldsymbol{u})=j} J_{\boldsymbol{u}}(d_0)^2 J_{\boldsymbol{u}}(P)^2, \tag{4.8}$$

$$A_{j+1,1}(D) = A_j(d_0) - A_{j,0}(D). \tag{4.9}$$

Theorem 1 gives the following corollary.

Corollary 1. *For an* $(sN, 2^n s)$*-design consisting of* s *subdesigns* $d_0, \ldots, d_{s-1}$ *defined as* (4.6),

(i) $A_{j,0}(D) = 0$ *if and only if for any* $J_{\boldsymbol{u}}(d_0) \neq 0$ *with* $wt(\boldsymbol{u}) = j$ *the corresponding* $J_{\boldsymbol{u}}(P) = 0$*;*

(ii) $A_{j+1,1}(D) = 0$ *if and only if for any* $J_{\boldsymbol{u}}(d_0) \neq 0$ *with* $wt(\boldsymbol{u}) = j$ *the corresponding* $J_{\boldsymbol{u}}(P) = \pm s$*.*

Clearly, $A_{1,0} = 0$ if and only if either any column of d_0 or that of P is balanced; $A_{2,1} = 0$ if and only if either column of d_0 is balanced or that of P is $(0, \ldots, 0)^T$ or $(1, \ldots, 1)^T$. Therefore, d_0 is set to a balanced design with $A_{1,0}(D) = A_{2,1}(D) = 0$.

Furthermore, considering the first five terms in (3.5), the initial design d_0 is set to a regular 2^{n-m} design with resolution $R \geq III$ and $J_{\boldsymbol{u}}(d_0) = 0$ for all $\boldsymbol{u}$'s with $wt(\boldsymbol{u}) = 1, 2$. Then $A_{1,0}(D) = A_{2,1}(D) = A_{2,0}(D) = A_{3,1}(D) = 0$. For the topic of regular fractional factorial designs, see Mukerjee and

Wu (2006) and Wu and Hamada (2021). Therefore, the SGMA criterion is equivalent to sequentially minimizing

$$(A_{3,0}, A_{4,1}, \ldots, A_{n,0}, A_{n+1,1}). \tag{4.10}$$

In the following constructions, we focus on the situation where d_0 is a given regular design of resolution $R \geq III$.

4.2 Construction of optimal PFD-based sliced factorial designs

$E(d_0, \mathcal{P})$ is used to denote the group of PFD-based SFDs generated by the regular factorial 2^{n-m} design d_0 and a switch matrix P from $\mathcal{P}$, where $\mathcal{P}$ contains all possible switch matrices. The switch matrix P is an $s \times n$ matrix, in which the elements $p_{0j} = 0$ for $j = 1, \ldots, n$ and $p_{ij} = 0, 1$ for $i = 1, \ldots, s-1$, $j = 1, \ldots, n$. Thus, an exhaustive search over all possible P configurations requires up to $O(2^{n(s-1)})$ operations. For example, considering a scenario with 10 factors and 5 platforms, an exhaustive search across all possible configurations of P would necessitate 2^{40} operations, which is impractical. To address this issue, we propose a systematic construction method for the switch matrix P to reduce computational complexity and generate the design of the optimal PFD-based $(sN, 2^n s)$-design among $E(d_0, \mathcal{P})$ for a given d_0.

Suppose that the defining contrast subgroup of d_0 is $\mathcal{G} = \{\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{2^m-1}\}$

with capacity 2^m. Without loss of generality, assume

(1) The vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ are linearly independent pencils and $\boldsymbol{u}_0$ is n-dimensional zero vector;

(2) Note that

$$U_{m\times n} = \begin{pmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_m \end{pmatrix} = \left(V_{m\times(n-m)},\ W^T_{m\times m}\right), \tag{4.11}$$

and the matrix $W_{m\times m}$ is invertible on $GF(2)$, where ^{T} represents the transposition of a matrix.

According to the (4.9) of Theorem 1, for a given regular d_0, sequentially minimizing SGWLP $(A_{3,0}, A_{4,1}, \ldots, A_{n,0}, A_{n+1,1}) = (A_{3,0}(D), A_3(d_0) - A_{3,0}(D), \ldots, A_{n,0}, A_n(d_0) - A_{n,0}(D))$, is equivalent to sequentially minimizing

$$(A_{3,0}(D), A_{4,0}(D), \ldots, A_{n,0}(D)).$$

According to the (4.8) of Theorem 1,

$$\begin{aligned} A_{j,0}(D) &= \frac{1}{(sN)^2}\sum_{wt(\boldsymbol{u})=j} J_{\boldsymbol{u}}(d_0)^2 J_{\boldsymbol{u}}(P)^2 \\ &= \frac{1}{(sN)^2}\sum_{\substack{wt(\boldsymbol{u})=j \\ J_{\boldsymbol{u}}(d_0)\neq 0}} J_{\boldsymbol{u}}(d_0)^2 J_{\boldsymbol{u}}(P)^2 \\ &= \frac{1}{s^2}\sum_{\substack{wt(\boldsymbol{u})=j \\ J_{\boldsymbol{u}}(d_0)\neq 0}} J_{\boldsymbol{u}}(P)^2 \\ &= \frac{1}{s^2}\sum_{\substack{wt(\boldsymbol{u})=j \\ \boldsymbol{u}\in\mathcal{G}}} J_{\boldsymbol{u}}(P)^2. \end{aligned}$$

Thus, sequentially minimizing the SGWLP is equivalent to sequentially minimizing $(B_3, \ldots, B_n)$, where $B_j = \sum_{\substack{wt(\boldsymbol{u})=j \\ \boldsymbol{u}\in\mathcal{G}}} J_{\boldsymbol{u}}(P)^2$.

Although there are $2^{n(s-1)}$ different switch matrices P, they may result in SFDs with the same SGWLP. Two SFDs whose corresponding switch matrices can be obtained from one another by a sequence of row permutations have the same SGWLP and are called equivalent. Consequently, a representative can be computed from each equivalence class. This leads to the following result.

Proposition 1. *For a given regular d_0, there are at most $\binom{2^m+s-2}{s-1}$ different sequences $(B_3, \ldots, B_n)$ determined by $J_{\boldsymbol{u}}(P)$ for all different switch matrices P.*

For a given $\boldsymbol{p}_{\boldsymbol{u}_1}, \ldots, \boldsymbol{p}_{\boldsymbol{u}_m}$,

$$P = \left(\mathbf{0}_{s\times(n-m)},\ (\boldsymbol{p}_{\boldsymbol{u}_1}, \ldots, \boldsymbol{p}_{\boldsymbol{u}_m})W^{-1}\right), \tag{4.12}$$

where W is defined in Equation (4.11) and $P\boldsymbol{u}_i^T = \boldsymbol{p}_{\boldsymbol{u}_i}, i = 1, \ldots, m$. The construction steps for the optimal $(sN, 2^n s)$-design among $E(d_0, \mathcal{P})$ are given in Algorithm 1. We provide two examples for this algorithm.

When $s \leqslant 2^m$, the optimal $(sN, 2^n s)$-design among $E(d_0, \mathcal{P})$ obtained by Algorithm 1 never includes repeats of the same flats, whereas when $s > 2^m$, although some of the flats must repeat, the maximum difference in the frequency of occurrence of each flat is at most 1. This observation is confirmed by Tables 1 and 2 in the Supplementary Material. To prove this result, it is necessary to consider not only the structure of the switch matrix P but also the defining contrast subgroup of d_0. We intend to closely investigate this topic in future work.

Algorithm 1

1: Given d_0 and the corresponding independent pencils $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$.

2: Given $\boldsymbol{p}_{\boldsymbol{u}_1}, \ldots, \boldsymbol{p}_{\boldsymbol{u}_m}$, obtain a sequence $(B_3, \ldots, B_n)$. Collect all $\binom{2^m+s-2}{s-1}$ sequences $(B_3, \ldots, B_n)$.

3: Choose the best sequence $(B_3, \ldots, B_n)$ from all $\binom{2^m+s-2}{s-1}$ sequences with the corresponding values of $\boldsymbol{p}_{\boldsymbol{u}_1}, \ldots, \boldsymbol{p}_{\boldsymbol{u}_m}$, denoted by $\boldsymbol{p}^{opt}_{\boldsymbol{u}_1}, \ldots, \boldsymbol{p}^{opt}_{\boldsymbol{u}_m}$.

4: Obtain the optimal switch matrix

$$P = \left(\mathbf{0}_{s\times(n-m)},\ (\boldsymbol{p}^{opt}_{\boldsymbol{u}_1}, \ldots, \boldsymbol{p}^{opt}_{\boldsymbol{u}_m})W^{-1}\right). \tag{4.13}$$

Ensure: The optimal SFD among $E(d_0, \mathcal{P})$ consists of s subdesigns $d_0, \ldots, d_{s-1}$ with d_i defined in (4.6), $i = 0, \ldots, s-1$.

Example 2. Consider the case of $(3 \cdot 8, 2^5 3)$-design. Let the initial design d_0 be a regular MA 2^{5-2} design with two independent pencils $\boldsymbol{u}_1 = (1,1,0,1,0)$, and $\boldsymbol{u}_2 = (1,0,1,0,1)$. The defining contrast subgroup is $\mathcal{G} = \{\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ with $\boldsymbol{u}_0 = (0,0,0,0,0)$ and $\boldsymbol{u}_3 = \boldsymbol{u}_1 + \boldsymbol{u}_2 = (0,1,1,1,1)$. Taking $\boldsymbol{p}_{\boldsymbol{u}_1} = (0,0,1)^T, \boldsymbol{p}_{\boldsymbol{u}_2} = (0,1,0)^T$ as an example, the switch matrix P can be obtained by (4.12). Then, $B_3 = J_{\boldsymbol{u}_1}(P)^2 + J_{\boldsymbol{u}_2}(P)^2 = 2$, $B_4 = J_{\boldsymbol{u}_3}(P)^2 = 4$, and $B_5 = 0$. Similarly, all eight sequences of (B_3, B_4, B_5) can be calculated with the best sequence being $(2,4,0)$ with respect to $\boldsymbol{p}_{\boldsymbol{u}_1}^{opt} = (0,0,1)^T$, $\boldsymbol{p}_{\boldsymbol{u}_2}^{opt} = (0,1,0)^T$. Thus, we obtain the optimal $(3 \cdot 8, 2^5 3)$-design among $E(d_0, \mathcal{P})$ given in Table 1 with the corresponding SGWLP $(0.22, 1.78, 0.11, 0.89, 0, 0)$.

Using the MA design d_0 for all platforms yields another design given in Table 2, denoted SMA.SS, where all subdesigns are the same as the regular 2^{5-2} MA design. The SGWLP of SMA.SS is $(2,0,1,0,0,0)$, and the constructed design in Table 1 is better.

Table 1: Optimal $(3{\cdot}8, 2^5 3)$-design among $E(d_0, \mathcal{P})$

	F_1	F_2	F_3	F_4	F_5	S
d_0	0	0	0	0	0	0
	0	0	1	0	1	0
	0	1	0	1	0	0
	0	1	1	1	1	0
	1	0	0	1	1	0
	1	0	1	1	0	0
	1	1	0	0	1	0
	1	1	1	0	0	0
d_1	0	0	0	0	1	1
	0	0	1	0	0	1
	0	1	0	1	1	1
	0	1	1	1	0	1
	1	0	0	1	0	1
	1	0	1	1	1	1
	1	1	0	0	0	1
	1	1	1	0	1	1
d_2	0	0	0	1	0	2
	0	0	1	1	1	2
	0	1	0	0	0	2
	0	1	1	0	1	2
	1	0	0	0	1	2
	1	0	1	0	0	2
	1	1	0	1	1	2
	1	1	1	1	0	2

Table 2: SMA.SS $(3{\cdot}8, 2^5 3)$-design

	F_1	F_2	F_3	F_4	F_5	S
d_0	0	0	0	0	0	0
	0	0	1	0	1	0
	0	1	0	1	0	0
	0	1	1	1	1	0
	1	0	0	1	1	0
	1	0	1	1	0	0
	1	1	0	0	1	0
	1	1	1	0	0	0
d_1	0	0	0	0	0	1
	0	0	1	0	1	1
	0	1	0	1	0	1
	0	1	1	1	1	1
	1	0	0	1	1	1
	1	0	1	1	0	1
	1	1	0	0	1	1
	1	1	1	0	0	1
d_2	0	0	0	0	0	2
	0	0	1	0	1	2
	0	1	0	1	0	2
	0	1	1	1	1	2
	1	0	0	1	1	2
	1	0	1	1	0	2
	1	1	0	0	1	2
	1	1	1	0	0	2

Algorithm 1 does not require d_0 to be an MA design. However, our search indicates that when the initial design is an MA design, the resulting optimal SFD among $E(d_0, \mathcal{P})$ tends to yield the best performance. The corresponding results are presented in Section S1 of the Supplementary Material. For further clarification, we provide Example 3.

Example 3. Consider the case of $(6 \cdot 16, 2^{10}6)$-design. Construct two SFDs D_1 and D_2. The initial designs of D_1 and D_2, denoted $d_0^{(1)}$ and $d_0^{(2)}$, respectively, are Design 10-6.1 and 10-6.3, respectively as listed in Table 3A.2 of Mukerjee and Wu (2006). Obtain the optimal $(6 \cdot 16, 2^{10}6)$-designs among $E(d_0^{(1)}, \mathcal{P})$ and $E(d_0^{(2)}, \mathcal{P})$ via Algorithm 1. The optimal switch matrices are as follows:

Table 3: Optimal switch matrices of D_1 and D_2

Optimal switch matrix of D_1										Optimal switch matrix of D_2									
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	1	0	1	0	1
0	0	0	0	1	1	0	0	1	1	0	0	0	0	1	1	1	0	0	1
0	0	0	0	1	0	1	1	0	1	0	0	0	0	1	0	1	1	1	0
0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	1	1	0
0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1

The SGWLPs of D_1 and D_2 are $(0, 8, 2, 16, 3.56, 12.44, \ldots)$ and $(0, 10, 1.67, 13.33, 3.56, 8.44, \ldots)$, respectively. Thus, D_1 outperforms D_2 according to the SGMA criterion.

Note that $d_0^{(1)}$ is an MA design, but as shown in Table 7 in Wang and Mee (2021), the 6-PFD with the initial design of $d_0^{(2)}$ is the best. Our conclusion differs because the SGWLP includes the term $A_{j,1}(D)$, which measures the overall aliasing between all j-factor $F_{i_1} \times \cdots \times F_{i_{j-1}} \times S$ and the general mean.

In general, when the SGMA criterion is considered, if the initial design d_0 is the design with less aberration, the corresponding optimal PFD-based $(sN, 2^n s)$-design among $E(d_0, \mathcal{P})$ is better with some exceptions. For example, consider Design 10-6.4, as listed in Table 3A.2 of Mukerjee and Wu (2006), as an initial design, denoted $d_0^{(3)}$. Although $d_0^{(2)}$ in Example 3 has less aberration than does $d_0^{(3)}$, the optimal PFD-based $(4 \cdot 16, 2^{10}4)$-designs among $E(d_0^{(3)}, \mathcal{P})$ is better than those among $E(d_0^{(2)}, \mathcal{P})$ based on the SGMA criterion, as shown in Table 2 of the Supplementary Material.

5. Sliced factorial designs with subdesigns of different run sizes

The previous section addresses the case where the run sizes of the subdesigns are equal. However, for some platforms, it might need to estimate additional effects. This situation requires increased degrees of freedom, which can be achieved by implementing multiple PFDs on those platforms. For instance, some platforms may require more observations than others. For this reason, not all factor combinations are feasible on every platform. For example, a smartwatch may not be able to accommodate two image-based factors and a brand logo on the same screen, whereas this combination is feasible on a desktop. Now revisit the example in Section 1. Let the three factors be labeled as 1, 2, and 3. Factors 1 and 2 (“Navigation” tab and “Login” tab) each have two levels, indicating whether or not all secondary menus are included, whereas factor 3 (“Research” tab) has two levels, representing the presence or absence of a hyperlink. Due to screen size limitation, a smartwatch cannot display all secondary menus and a hyperlink at the same time, whereas a desktop can accommodate this. Therefore, if the sample size is sufficient, all eight versions can be tested on the desktop platform. On the smartwatch, however, a fractional factorial design with generator $C = AB + \mathbf{1}$ can be used to test a subset of these combinations.

To solve the above problem, we explore the case where a platform may consist of multiple flats. Here the initial design may differ from any individual subdesign. Let d_0^* represent the initial design to distinguish it from the previous section. For $i = 0, \ldots, s-1$, each subdesign d_i is either isomorphic with respect to the initial design d_0^* or contains multiple such isomorphic designs. Obviously, the run size of d_i is an integral multiple of that of d_0^*.

Let $v_i \geqslant 1$, $i = 0, \ldots, s-1$, represent the number of isomorphic designs of d_0^* in d_i with $v_0 + v_1 + \cdots + v_{s-1} = q$. SFD is a $(qN, 2^n s)$-design, where N is the run size of d_0^*.

Similar to Section 4.1, the jth column of subdesign d_i is defined as

$$\boldsymbol{d}_{ij} = \boldsymbol{p}_{ij} \oplus \boldsymbol{d}_{0j}^*, \tag{5.14}$$

where $\boldsymbol{p}_{ij} = (p_{ij}^{(1)}, \ldots, p_{ij}^{(v_i)})^T$, $p_{ij}^{(h)} = 0, 1$ for $i = 0, \ldots, s-1$, $j = 1, \ldots, n$ and $h = 1, \ldots, v_i$. Here, without loss of generality, let $p_{0j}^{(1)} = 0$ for all $j = 1, \ldots, n$. For convenience, let $P_i = (\boldsymbol{p}_{i1}, \ldots, \boldsymbol{p}_{in})$ for $i = 0, \ldots, s-1$, and combine them together to obtain a switch matrix $P = (P_0^T, P_1^T, \ldots, P_{s-1}^T)^T$. Furthermore, according to the definition of J - characteristics, $J_{\boldsymbol{u}}(P) = \sum_{i=0}^{s-1} J_{\boldsymbol{u}}(P_i)$ and $J_{\boldsymbol{u}}(P_i) = \sum_{\boldsymbol{x} \in P_i} (-1)^{\langle \boldsymbol{u}, \boldsymbol{x} \rangle}$, where $\boldsymbol{x}$ is a row vector of the submatrix P_i. Then, we obtain the following lemma.

Lemma 1. *The J-characteristic of subdesign d_i can be expressed as*

$$J_{\boldsymbol{u}}(d_i) = J_{\boldsymbol{u}}(d_0^*) J_{\boldsymbol{u}}(P_i), \quad i = 0, \ldots, s-1. \tag{5.15}$$

Lemma 1 gives the following theorem. This shows that, even if the run sizes of the subdesigns are different, the SGWLP of an SFD is still determined by the initial design d_0^* and the switch matrix P. The proof of Theorem 2 is similar to that of Theorem 1 and thus omitted.

Theorem 2. *For a* $(qN, 2^n s)$*-design consisting of* s *subdesigns* $d_0, \ldots, d_{s-1}$ *defined in* (5.14),

$$A_{j,0}(D) = \frac{1}{(qN)^2} \sum_{wt(\boldsymbol{u})=j} J_{\boldsymbol{u}}(d_0^*)^2 J_{\boldsymbol{u}}(P)^2, \tag{5.16}$$

$$A_{j+1,1}(D) = \big(s/q\big)^2 A_j(d_0^*) - A_{j,0}(D). \tag{5.17}$$

From Theorem 2, we obtain the following corollary.

Corollary 2. *For a* $(qN, 2^n s)$*-design consisting of* s *subdesigns* $d_0, \ldots, d_{s-1}$ *defined as* (5.14),

(i) $A_{j,0}(D) = 0$ *if and only if for any* $J_{\boldsymbol{u}}(d_0^*) \neq 0$ *with* $wt(\boldsymbol{u}) = j$ *the corresponding* $J_{\boldsymbol{u}}(P) = 0$*;*

(ii) $A_{j+1,1}(D) = 0$ *if and only if* $J_{\boldsymbol{u}}(d_0^*) = 0$ *with* $wt(\boldsymbol{u}) = j$*;*

(iii) If $J_{\boldsymbol{u}}(d_0^*) \neq 0$ *with* $wt(\boldsymbol{u}) = j$ *exists, the minimum value of* $A_{j+1,1}(D)$ *is* $\big((s/q)^2 - 1\big)A_j(d_0^*)$ *if and only if for any* $J_{\boldsymbol{u}}(d_0^*) \neq 0$ *with* $wt(\boldsymbol{u}) = j$ *the corresponding* $J_{\boldsymbol{u}}(P) = \pm q$*.*

Consider the first five terms in (3.5). The values of v_i, $i = 0, \ldots, s-1$ are given and are not all one, implying that the slice factor is not balanced. Thus, $A_{1,1}$ is a positive constant. We can select the initial design d_0^* to be a regular 2^{n-m} design with resolution $R \geq III$; then $J_{\boldsymbol{u}}(d_0^*) = 0$ for all $\boldsymbol{u}$s with $wt(\boldsymbol{u}) = 1, 2$, and $A_{1,0}(D) = A_{2,1}(D) = A_{2,0}(D) = A_{3,1}(D) = 0$.

In this case, the SGMA criterion is equivalent to sequentially minimizing (4.10).

Example 4. Consider the case of an $(6 \cdot 16, 2^6 3)$-design, where the run sizes of d_0, d_1 and d_2 are 16, 32 and 48, respectively. The initial design d_0^* is Design 6-2.1 as listed in Table 3A.2 of Mukerjee and Wu (2006).

Using Algorithm 1, we can obtain three nonisomorphic optimal $(6 \cdot 16, 2^6 3)$-designs, D_1, D_2 and D_3, among $E(d_0^*, \mathcal{P})$, in which the SGWLPs are all the same. The switch matrices of D_1, D_2 and D_3 are listed as below.

Table 4: Optimal switch matrices of D_1, D_2 and D_3

	Optimal switch matrix of D_1						Optimal switch matrix of D_2						Optimal switch matrix of D_3					
P_0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P_1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	1	1
	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0
P_2	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	1
	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

On the basis of the SGMA criterion, D_1, D_2 and D_3 are optimal $(6 \cdot 16, 2^6 3)$-designs among $E(d_0^*, \mathcal{P})$. However, their subdesigns may not be optimal. To maintain consistency with Table 4, let the subdesigns of D_1 be $d_0^{(1)}$, $d_1^{(1)}$, and $d_2^{(1)}$, those of D_2 be $d_0^{(2)}$, $d_1^{(2)}$, and $d_2^{(2)}$; and those of D_3 be $d_0^{(3)}$, $d_1^{(3)}$, and $d_2^{(3)}$. Table 4 indicates that $d_1^{(1)}$ and $d_1^{(3)}$ are fold-over designs with less aberration than $d_1^{(2)}$ has. Similarly, $d_2^{(2)}$ and $d_2^{(3)}$ are three-quarter fractional designs with less aberration than $d_2^{(1)}$ has. Therefore, D_3 is preferable to D_1 and D_2 in practice.

This example demonstrates that row permutations of the switch matrix P produce SFDs with the same SGWLP, but their subdesigns may exhibit distinct generalized wordlength patterns. To further distinguish the SFDs, we propose a secondary criterion as recommended by one reviewer: summing the generalized wordlength patterns of the separate slices. Using this approach, the summed A_4 values for D_1, D_2, and D_3 are 5.22, 6.33, and 4.33, respectively. Clearly, D_3 is the best of these three optimal $(6 \cdot 16, 2^6 3)$-designs among $E(d_0^*, \mathcal{P})$. Specifically, we recommend using d_i as a fold-over design when the run size of the subdesign d_i is twice that of the initial design d_0^*, in which case d_i achieves the lowest GWLP. Similarly, we recommend selecting d_i as a three-quarter design if the run size of d_i is three times that of d_0^*. For details on fold-over and three-quarter designs, see Mee (2009).

6. Concluding remarks

Sliced factorial designs are useful for online experiments conducted across multiple platforms. The construction of such designs for any number of platforms is an urgent problem in practice. To address this modern challenge, we provide a sliced generalized wordlength pattern based on the sliced effect hierarchy principle to assess the overall aliasing between all effects. We presented the sliced factorial designs for online multi-platform experiments, where the interactions between the slice factor and the design factors are not only nonnegligible but also important. Sliced factorial designs equally apply to physical experiments, such as experiments run in parallel on a collection of bioreactors or parallel production lines in manufacturing. For information on parallel bioreactors, see Akgün et al. (2004), Gill et al. (2008), Bareither and Pollard (2011), Mandenius (2016); for parallel production lines, see Burman (1995), Nahas et al. (2009), Verbiest et al. (2019), and Xi et al. (2022).

Owing to the numerous potential SFD candidates for certain parameter settings, conducting an exhaustive search for the optimal SFD is impractical. Therefore, we focused only on constructing the PFD-based SFDs, and provided some theoretical results on their characteristics, as well as a construction method for obtaining optimal PFD-based SFDs. We discuss the cases where the run sizes of all slices are either equal or unequal in Sections 4 and 5, respectively.

For online experiments, the sample size on each platform depends on the traffic flow to the page, resulting in unbalanced sample sizes. In practice, experimental designs are typically formulated before users are assigned to different versions, and the exact sample size per platform is often unknown. In such cases, an unequal variance model or a random effects model can be used for the analysis. Certainly, as the referees pointed out, when prior information on sample sizes is available, the design criterion can be further refined to account for unbalanced data, potentially leading to more efficient designs. We have identified this extension as a promising direction for future research.

In practice, if the numbers m and s are not too large, we can obtain the optimal PFD-based SFDs by Algorithm 1. We briefly mention an idea for constructing optimal PFD-based SFDs when m and s are relatively large. For $\boldsymbol{u} \in \mathcal{G}$, if the numbers of 0s and 1s are the same or differ by one, then $J_{\boldsymbol{u}}(P)$ is minimized. We construct a specific switch matrix P, incorporating $\boldsymbol{p}_{\boldsymbol{u}}$, such that the sequence $(B_3, \ldots, B_m)$ can be minimized sequentially. We plan to further study this construction in a follow-up project.

Supplementary Materials

The online supplementary materials provide optimal $(sN, 2^n s)$-designs among $E(d_0, \mathcal{P})$ and proofs of the theoretical results.

Acknowledgments

Han, Wang and Sun are supported by the National Natural Science Foundation of China (No. 12371259 and 11971098), the Fundamental Research Funds for the Central Universities, China (No. 2412023YQ003), and the National Key Research and Development Program of China (No. 2020YFA 0714102 and 2022YFA1003701). Peter Chien and Fasheng Sun are the corresponding authors. We would like to thank the editor, associate editor and two referees for their useful comments, which have significantly improved the article.

References

Akgün, A., B. Maier, D. Preis, B. Roth, R. Klingelhöfer, and J. Büchs (2004). "A novel parallel shaken bioreactor system for continuous operation". *Biotechnology progress 20*(6), 1718–1724.

Bareither, R. and D. Pollard (2011). "A review of advanced small-scale parallel bioreactor technology for accelerated process development: Current state and future need". *Biotechnology progress 27*(1), 2–14.

Burman, M. H. (1995). *"New results in flow line analysis"*. Ph. D. thesis, Massachusetts Institute of Technology.

Cheng, C.-S. (2014). *"Theory of factorial design"*. Chapman and Hall/CRC Boca Raton, FL, USA.

Connor, W. S. and S. Young (1961). *"Fractional factorial designs for experiments with factors at two and three levels"*. US Government Printing Office.

Edwards, D. J. and R. W. Mee (2023). "Structure of nonregular two-level designs". *Journal of the American Statistical Association 118*(542), 1222–1233.

Gill, N. K., M. Appleton, F. Baganz, and G. J. Lye (2008). "Design and characterisation of a miniature stirred bioreactor system for parallel microbial fermentations". *Biochemical engineering journal 39*(1), 164–176.

Guo, Q., X. Deng, and P. Chien (2024). "Multi-layer sliced design and analysis with application to ai assurance". *Technometrics* under review.

Haizler, T. and D. M. Steinberg (2021). "Factorial designs for online experiments". *Technometrics 63*(1), 1–12.

Kong, X., M. Ai, and K. L. Tsui (2018). "Flexible sliced designs for computer experiments". *Annals of the Institute of Statistical Mathematics 70*, 631–646.

Mandenius, C.-F. (2016). *"Bioreactors: design, operation and novel applications"*. John Wiley & Sons.

Mee, R. W. (2009). *"A comprehensive guide to factorial two-level experimentation"*. Springer.

Mukerjee, R. and C. F. J. Wu (2006). *"A modern theory of factorial design"*. Springer Science & Business Media.

Nahas, N., M. Nourelfath, and D. Ait-Kadi (2009). "Selecting machines and buffers in unreliable series-parallel production lines". *International Journal of Production Research 47*(14), 3741–3774.

Qian, P. Z. G. (2012). "Sliced latin hypercube designs". *Journal of the American Statistical Association 107*(497), 393–399.

Qian, P. Z. G. and C. F. J. Wu (2009). "Sliced space-filling designs". *Biometrika 96*(4), 945–956.

Sadeghi, S., P. Chien, and N. Arora (2020). "Sliced designs for multi-platform online experiments". *Technometrics 62*(3), 387–402.

Sadeghi, S., T. Hung, P. Chien, and N. Arora (2024). "A sliced design approach for conducting an industry email campaign with four platforms". *New England Journal of Data Science 2*(3), 311–322.

Tang, B. and L. Y. Deng (1999). "Minimum g_2-aberration for nonregular fractional factorial designs". *The Annals of Statistics 27*(6), 1914–1926.

Verbiest, F., T. Cornelissens, and J. Springael (2019). "A matheuristic approach for the design of multiproduct batch plants with parallel production lines". *European Journal of Operational Research 273*(3), 933–947.

Wang, C. and R. W. Mee (2021). "Two-level parallel flats designs". *The Annals of Statistics 49*(5), 3015–3042.

Wu, C. F. J. and M. S. Hamada (2021). *"Experiments: planning, analysis, and optimization"*.

John Wiley & Sons.

Wu, C. F. J. and H. Xu (2001). "Generalized minimum aberration for asymmetrical fractional factorial designs". *The Annals of Statistics 29*(4), 1066–1077.

Xi, S., J. M. Smith, Q. Chen, N. Mao, H. Zhang, and A. Yu (2022). "Simultaneous machine selection and buffer allocation in large unbalanced series-parallel production lines". *International Journal of Production Research 60*(7), 2103–2125.

Yang, J., M. Q. Liu, and D. K. J. Lin (2014). "Construction of nested orthogonal latin hypercube designs". *Statistica sinica 24*(1), 211–219.

Zhang, R. and D. Park (2000). "Optimal blocking of two-level fractional factorial designs". *Journal of Statistical Planning and Inference 91*(1), 107–121.

KLAS and School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin 130024, China

E-mail: hanzj087@nenu.edu.cn

School of Statistics, Jilin University of Finance and Economics, Changchun, Jilin 130117, China

E-mail: wangdy798@126.com

KLAS and School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin 130024, China

E-mail: sunfs359@nenu.edu.cn

Department of Statistics, University of Wisconsin-Madison, Madison, WI 53706, USA

E-mail: peter.chien@wisc.edu